علم اللُّغة والتِّقنيات المُعاصِرة

د.المُعتزّ بالله السَّعيد

moataz@cu.edu.eg

مرحبا بكم

welcome

5 November 2014

المُحاضَرة السَّابعة

في المُدَوَّنات اللُّغَوِيَّة

٣

5 November 2014

- عَنونة المدَوَّنات اللُّغَوِيَّة -

5 November 2014

- عَنونة المدَوَّنات اللُّغَوِيَّة -

5 November 2014

عَنونة المدَوَّنات اللُّغَوِيَّة

- عنونة المدَوَّنة اللُّغَوِيَّة (Corpus Annotation) عمليَّةٌ وسيطةٌ، تنتقلُ خلالَها النُّصُوصُ من صُورتها الأوَّليَّة (الخام) إلى صُورةٍ يسهُلُ التَّعامُلُ معها آليًّا (مُعَنونة). ويُقصَدُ بعنونة نُصُوص المدَوَّنات اللُّغَوِيَّة إضافة معلوماتٍ توضِيحيَّة توصِيفِيَّة لكُلِّ وحدة لُغَوِيَّةٍ [كلمة] في النُّصُوص على حدة، بحيثُ تُصبِحُ المعلوماتُ مُلازِمةً للوحدات.
- تُمَهِّدُ هذه العمليَّةُ للمعالَجة الآليَّة لنُصُوص المدَوَّنات اللُّغَوِيَّة، كما تُساعد في إجراء الاختبارات اليدَويَّة للنُّصُوص، حيثُ تُقَلِّلُ من الجهد المبذول في مُتابعة وإحصاء توصِيفات المفرَدات.

عَنونة المدَوَّنات اللُّغَوِيَّة

- يعمدُ بعضُ صُنَّاع المدَوَّنات اللُّغَوِيَّة المستخدمة في صناعة المعجم إلى الإبقاء على مُدَوَّناتهم في صُورتها الخام، حرصًا على تماسُك هيئة النَّصُوص عند الاستشهاد بها أو البحث فيها، بينما يعمدُ غيرُهم إلى وضع المدَوَّنات اللُّغَوِيَّة المعجميَّة في صُورَتَين، إحداهما خام (للاستشهاد والبحث) والأخرى مُعَنونة (للإحصاء والمعالَجة الآليَّة).
- تتَضَمَّنُ عَنونات المدَوَّنات اللُّغَوِيَّة معلوماتٍ توصِيفِيَّة لُغَوِيَّة، وأخرى غير لُغَوِيَّة، تتمثَّلُ في المعلومات الببليوجرافيَّة، والخصائص الشَّكليَّة للنُّصُوص (مثل: نوع الخُطُوط وأحجامها [المُدَوَّنات المكتوبة]، ومواضع النَّبر والتَّنغيم والمقاطع الصَّوتِيَّة [المُدوَّنات المنطوقة]).

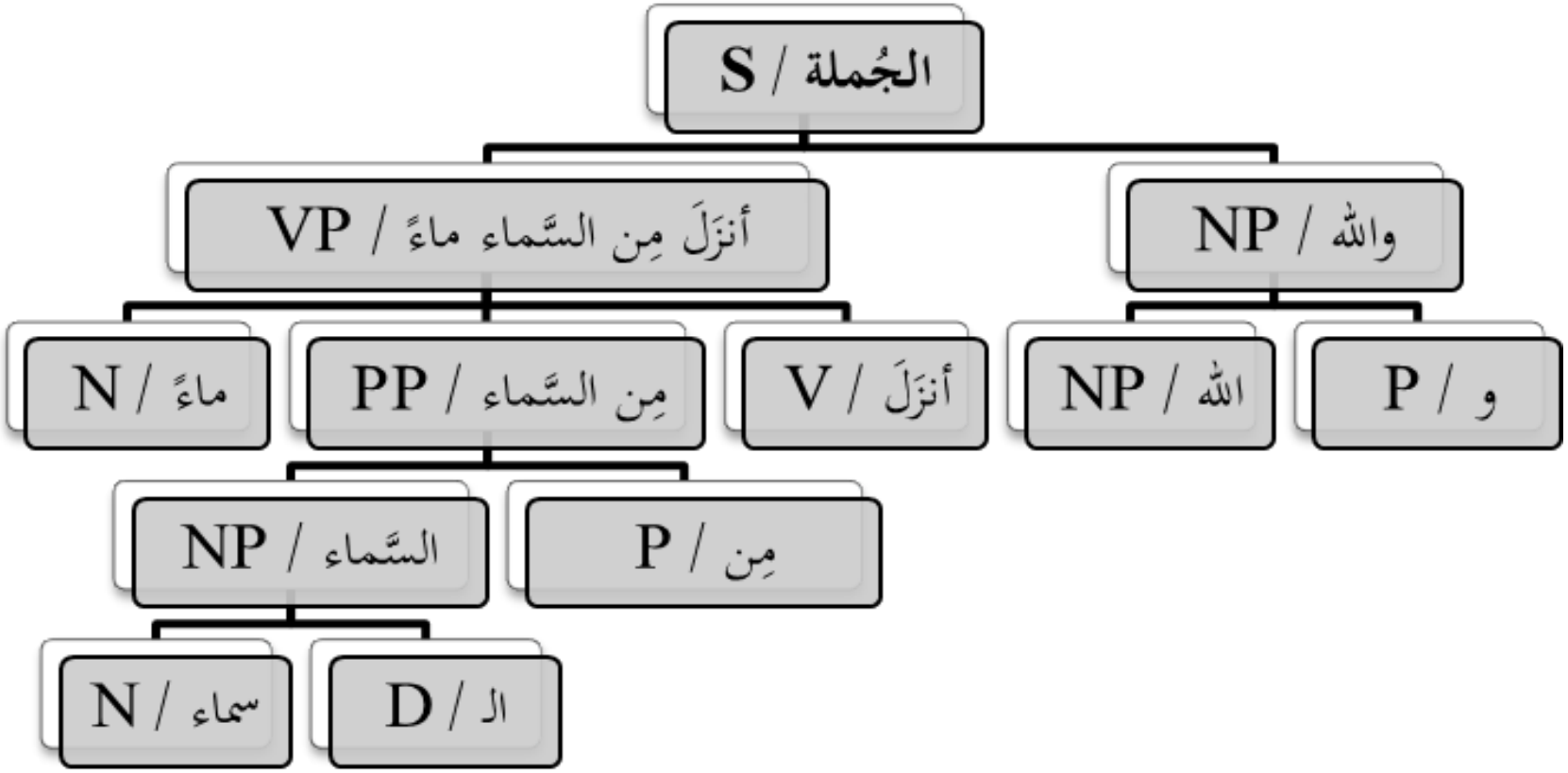
عَنونة المدَوَّنات اللُّغَوِيَّة

- العَنونة التَّركيبيَّة - العَنونة بتعيين أقسام الكلام:

النَّصُّ قبلَ العَنوَنة
الآن.. ما الذي يشغل المثقفين المصريين أو العرب؟
النَّصُّ بعدَ العَنوَنة
الآن RB / .. / ما RP / الذي WP / يشغل VBP / المثقفين NNS / المصريين NNS / أو CC / العرب NNS / ؟ /
رُمُوز عَنونة أقسام الكلام PoS Tags
الظَّرف RB، الأداة RP، الاسم الموصول WP، الفِعل المضارع VBP، الأسماء الشَّائعة (للجمع) NNS، حرف العطف CC، الضَّمير PRP، الأسماء الشَّائعة (للمفرَد) NN

عَنونة المدَوَّنات اللُّغَوِيَّة

- العَنونة التَّركيبيَّة - العَنونة بتحليل أقسام الكلام:



عَنونة المدَوَّنات اللُّغَوِيَّة

- **العَنونة الدِّلاليَّة ١:**

السِّياق			الدِّلالة
فرفعَ البلقيطيّ	حاجبيه	الخفيفين .. وقال: .. أنعِم وأكرِم	١
وتساءَلَ وهُو يَنتِفُ	حاجبه	الأيسر متى يمتلئ جيبي بنقود الحُكومة	١
فدخَلَ عليه	حاجبه	فأخبرَه أنَّ رَجُلا من الخوارج جيءَ به	٢
فلما طالَ انتظارُه، قالَ	للحاجب	إن أذِنتَ لي عليه، وإلاّ هجَوتُ اليَمَن .	٢
وقالَ جمال الدِّين بن	الحاجب	: ابن عبد السَّلام أفقهُ من الغزالي	٣
ومِن الخُطَباء عُطارد بن	حاجب	بن زرارة، وقد خَطَبَ أمام الرَّسُول	٣
(١) الشَّعْر النَّابِت فَوقَ العَيْن، (٢) خازن الباب وحارِسه، (٣) عَلَم / من أسماء الذُّكور			

عَنونة المدَوَّنات اللُّغَوِيَّة

- العَنونة الدِّلاليَّة ٢:

PoS	السِّياق (من: القُرآن الكريم)	(م / د / PoS)
N	{وَظَلَّلْنَا عَلَيْكُمُ الْغَمَامَ وَأَنزَلْنَا عَلَيْكُمُ المنَّ وَالسَّلْوَى}	(.. / ١ / ١)
N	{يَا أَيُّهَا الَّذِينَ آمَنُواْ لاَ تُبْطِلُواْ صَدَقَاتِكُم بِالمنِّ}	(.. / ٢ / ١)
N	{قَالُوا يَا وَيْلَنَا مَن بَعَثَنَا مِن مَّرْقَدِنَا}	(.. / ٣ / ١)
V	{وَنُرِيدُ أَن نَّمُنَّ عَلَى الَّذِينَ اسْتُضْعِفُوا فِي الْأَرْضِ}	(.. / ٤ / ٢)
V	{يَمُنُّونَ عَلَيْكَ أَنْ أَسْلموا قُل لَّا تَمُنُّوا عَلَيَّ إِسْلَامَكُم}	(.. / ٢ / ٢)
P	{فَتَلَقَّى آدَمُ مِن رَّبِّهِ كَلماتٍ فَتَابَ عَلَيْهِ}	(.. / ٥ / ٣)

عَنونة المدَوَّنات اللُّغَوِيَّة

- التَّرميز:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
    <o:Title>اللائحة الأساسية لمجلس النواب المصري</o:Title>
    <o:Author>المُشَرِّع</o:Author>
    <o:Created>2009-08-14T12:18:00Z</o:Created>
    <o:LastSaved>2009-08-14T12:18:00Z</o:LastSaved>
    <resp>compiled by</resp>    <o:Words>1565</o:Words>
    <o:Characters>8921</o:Characters>
    <o:Lines>74</o:Lines>
    <name>Almo3taz Bellah</name>
    <w:t>اللائحة الأساسية لمجلس النواب المصري</w:t>
    <w:t>اللائحة الأساسية التي وافق عليها مجلس النواب المصري وصدر بها الأمر العالي</w:t>
    <w:t>1882 فبراير سنة 7  (في 18 ربيع الأول سنة 1299 هـ)</w:t>
    <w:t>نحن خديو مصر</w:t>
    <w:t>1881 بعد الاطلاع على أمرنا الصادر بتاريخ 11 ذي القعدة سنة 1298 الموافق 4 أكتوبر سنة .</w:t>
    <w:t>وبناء على ما قرره مجلس النواب، وموافقة رأي مجلس نظارنا .</w:t>
```

شُكرًا لكم

5 November 2014

للتَّواصُل ..

- د. المُعتزّ بالله السَّعيد

moataz@cu.edu.eg

almoataz2006@skype.com

almo3tazbellah@facebook.com

www.ask.fm/AlmoatazBAlSaid

http://scholar.cu.edu.eg/moataz

http://cairo.academia.edu/AlmoatazBAlSaid